

NEWS | 01 March 2018

Colossal family tree reveals environment's influence on lifespan

Genetics explains only a small part of differences in how long a person lives, finds analysis that links 13 million people.

[Erika Check Hayden](#)



The world's largest family tree shows ancestry links over hundreds of years. Credit: Ulrich Doering/Alamy

Did you forget your mother's birthday this year? Brace yourself: your family tree may

now include the birthdays of 13 million people.

Computational biologist Yaniv Erlich of Columbia University in New York City and his colleagues have used crowdsourced data to make a family tree that links 13 million people. The ancestry chart, described today in *Science*¹, is believed to be the largest verified resource of its kind – spanning an average of 11 generations.

Erlich's team analysed the birth and death dates of the people in this tree, and calculated whether individuals were more likely to have died at similar ages if they were closely related. The group concludes that heredity explains only about 16% of [the difference in lifespans for these individuals](#). Most of the differences were down to other factors, such as where and how people lived.

“This is a real tour de force,” says genetic epidemiologist Braxton Mitchell of the University of Maryland School of Medicine in Baltimore. “This is a great example of using large, publicly available data sets to do interesting research.”

Live long and prosper

Scientists already suspected that environment has more influence than genes on how long people live. But Erlich estimates that genes have even less of a role than researchers had thought.

Some studies, such as one published by Mitchell's group in 2001², have estimated that genes determine about one-quarter of the variation in people's lifespan.

Erlich's finding proves the power of [extremely large family trees, or genealogies](#), says Lisa Cannon-Albright, a geneticist at the University of Utah School of Medicine in Salt Lake City.

“These kinds of resources will be a powerful piece of future genetics research,” she

says.

Erlich says that “good” genes might extend a person’s life by an average of five years. Some environmental factors make a much bigger impact on longevity; smoking, for instance, can subtract ten years.

Geneticists have long used family trees to study how genetics influence many traits, such as disease risk. But it can be costly and difficult to assemble databases of family records that contain vast numbers of people. Erlich’s study is one of many under way that are now assembling digital records into very large family trees^{3,4}; some have identified genes linked to illnesses such as cancer and Alzheimer’s disease⁵.



This 6,000-person family tree was cleaned and organized using graph theory. Individuals are shown in green, spanning seven generations; marriages are depicted in red. Credit: Columbia University

Data deluge

Erlich's study used data from an online genealogy tool, Geni.com. He is the chief scientific officer of MyHeritage, Geni's parent company, in Or Yehuda, Israel.

The analysis drew on data on roughly 86 million people whose records were uploaded by Geni users. That's an order of magnitude more participants than are included in [the largest consumer genetic-testing database](#).

"The sheer number of participants is crazy," says computational genomicist Atul Butte of the University of California, San Francisco. "You can only get data sets like this with crowdsourcing. It's really impressive."

Erlich's team used the data to analyse the migration and marriage patterns of people listed on Geni. For instance, before 1750, the researchers found, most Americans and Europeans in the database married someone who lived at most 10 kilometres from their birthplace. By 1950, most Americans and Europeans had to travel at least 100 kilometres from their home towns to find a spouse..

In other words, your parents probably travelled farther than any of their ancestors to start your family. The least you can do is remember their birthdays.

doi: <https://doi.org/10.1038/d41586-018-02596-9>

References

1. Kaplanis, J. *et al. Science* <http://dx.doi.org/10.1126/science.aam9309> (2018).

2. Mitchell, B. D. *et al. Am. J. Med. Genet.* **102**, 346–352 (2001).

3. Polubriaginof, F. *et al.* Preprint on bioRxiv at <http://dx.doi.org/10.1101/066068> (2016).

4. Cannon-Albright, L. A. *et al. Genet. Med.* **15**, 541–547 (2013).

5. Ridge, P. G. *et al. Genome Med.* **9**, 100 (2017).
